Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways

Georg Lenz[a], George W. Wright[b], N. C. Tolga Emre[a], Holger Kohlhammer[a], Sandeep S. Dave[a], R. Eric Davis[a], Shannon Carty[a], Lloyd T. Lam[a], A. L. Shaffer[a], Wenming Xiao[c], John Powell[c], Andreas Rosenwald[d], German Ott[d,e], Hans Konrad Muller-Hermelink[d], Randy D. Gascoyne[f], Joseph M. Connors[f], Elias Campo[g], Elaine S. Jaffe[h], Jan Delabie[i], Erlend B. Smeland[j,k], Lisa M. Rimsza[l], Richard I. Fisher[m,n], Dennis D. Weisenburger[o], Wing C. Chan[o], and Louis M. Staudt[a,q]

[a]Metabolism Branch, [b]Biometric Research Branch, [c]Center for Information Technology, and [h]Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892; [d]Department of Pathology, University of Würzburg, 97080 Würzburg, Germany; [e]Department of Clinical Pathology, Robert-Bosch-Krankenhaus, 70376 Stuttgart, Germany; [f]British Columbia Cancer Agency, Vancouver, British Columbia, Canada V5Z 4E6; [g]Hospital Clinic, University of Barcelona, 08036 Barcelona, Spain; [i]Pathology Clinic and [j]Institute for Cancer Research, Rikshospitalet Hospital, Oslo, Norway; [k]Centre for Cancer Biomedicine, Faculty Division the Norwegian Radium Hospital, N-0310, Oslo, Norway; [l]Department of Pathology, University of Arizona, Tucson, AZ 85724; [m]Southwest Oncology Group, 24 Frank Lloyd Wright Drive, Ann Arbor, MI 48106 ; [n]James P. Wilmot Cancer Center, University of Rochester, Rochester, NY 14642; and [o]Departments of Pathology and Microbiology, University of Nebraska, Omaha, NE 68198

Edited by Ira Pastan, National Cancer Institute, National Institutes of Health, Bethesda, MD, and approved July 7, 2008 (received for review May 2, 2008)

Gene-expression profiling has been used to define 3 molecular subtypes of diffuse large B-cell lymphoma (DLBCL), termed germinal center B-cell-like (GCB) DLBCL, activated B-cell-like (ABC) DLBCL, and primary mediastinal B-cell lymphoma (PMBL). To investigate whether these DLBCL subtypes arise by distinct pathogenetic mechanisms, we analyzed 203 DLBCL biopsy samples by high-resolution, genome-wide copy number analysis coupled with gene-expression profiling. Of 272 recurrent chromosomal aberrations that were associated with gene-expression alterations, 30 were used differentially by the DLBCL subtypes ($P < 0.006$). An amplicon on chromosome 19 was detected in 26% of ABC DLBCLs but in only 3% of GCB DLBCLs and PMBLs. A highly up-regulated gene in this amplicon was *SPIB*, which encodes an ETS family transcription factor. Knockdown of *SPIB* by RNA interference was toxic to ABC DLBCL cell lines but not to GCB DLBCL, PMBL, or myeloma cell lines, strongly implicating *SPIB* as an oncogene involved in the pathogenesis of ABC DLBCL. Deletion of the *INK4a*/*ARF* tumor suppressor locus and trisomy 3 also occurred almost exclusively in ABC DLBCLs and was associated with inferior outcome within this subtype. *FOXP1* emerged as a potential oncogene in ABC DLBCL that was up-regulated by trisomy 3 and by more focal high-level amplifications. In GCB DLBCL, amplification of the oncogenic mir-17–92 microRNA cluster and deletion of the tumor suppressor *PTEN* were recurrent, but these events did not occur in ABC DLBCL. Together, these data provide genetic evidence that the DLBCL subtypes are distinct diseases that use different oncogenic pathways.

gene-expression profiling | oncogenes | tumor suppressor genes | comparative genomic hybridization

Diffuse large B-cell lymphoma (DLBCL) is the most common type of non-Hodgkin's lymphoma (1). DLBCL can be cured using anthracycline-based chemotherapy regimens in only 40%–50% of patients, suggesting that DLBCL is a heterogeneous diagnostic category (2). Indeed, gene-expression profiling studies have distinguished 3 molecular subtypes of DLBCL known as "germinal center B-cell-like" (GCB) DLBCL, "activated B-cell-like" (ABC) DLBCL, and "primary mediastinal B-cell lymphoma" (PMBL) (3–7). GCB DLBCLs seem to arise from normal germinal center B cells, whereas ABC DLBCLs may arise from postgerminal center B cells that are arrested during plasmacytic differentiation, and PMBLs may arise from thymic B cells (8). Patients with these DLBCL subtypes have significantly different survival rates following chemotherapy (3–6), leading to the current proposal that they represent distinct biological disease entities (8).

Recent studies using fluorescence *in situ* hybridization, conventional comparative genomic hybridization (CGH), or low-resolution array-based CGH (aCGH) have shown that certain genetic aberrations occur at different frequencies among DLBCL subtypes (4, 9–11). The t(14, 18) translocation involving *BCL2* and amplification of *REL* was detected in 45% and 16% of GCB DLBCLs, respectively, but never in ABC DLBCLs (4). Roughly one-quarter of ABC DLBCLs had trisomy 3 or gain/amplification of chromosome arm 3q, but these abnormalities never were detected in GCB DLBCLs (9, 10). ABC DLBCLs were characterized further by frequent gain of 18q and loss of 6q (9, 10).

These relatively low-resolution techniques were unable to pinpoint the presumed cancer-relevant target gene in most instances. To overcome these limitations, we combined aCGH using high-density whole-genome microarrays with gene expression profiling. By this concerted approach, we identifiedchromosomal aberrations that were significantly more frequent in a particular DLBCL subtype than in the others, and some of these aberrations were associated with clinical outcome. This analysis revealed oncogenic pathways that are used differentially by the DLBCL subtypes, reinforcing the view that they represent pathogenetically distinct diseases.

Results and Discussion

Definition of Recurrently Altered Minimal Common Regions (MCRs). aCGH was performed on 203 DLBCL samples, including 74 ABC DLBCLs, 72 GCB DLBCLs, 31 PMBLs, and 26 unclassified DLBCLs from a previously studied cohort of patients (4). We used oligonucleotide microarrays covering the entire genome at 5-kb average spacing, allowing us to delineate copy number changes with great precision. To identify chromosomal aberrations associated with gene-expression changes, we profiled gene expression in the same samples and developed a statistical algorithm termed "Gene Expression and Dosage Integrator" (GEDI) to merge the data sets (Fig. 1). For each case, GEDI divides the aCGH data into intervals with statistically similar gene dosage, and each interval is classified as an amplification,

Author contributions: G.L. and L.M.S. designed research; G.L., N.C.T.E., H.K., S.S.D., R.E.D., S.C., L.T.L., and A.L.S. performed research; A.R., G.O., H.K.M.-H., R.D.G., J.M.C., E.C., E.S.J., J.D., E.B.S., L.M.R., R.I.F., D.D.W., and W.C.C. contributed new reagents; G.L., G.W.W., W.X., J.P., and L.M.S. analyzed data; and G.L., G.W.W., and L.M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[q]To whom correspondence should be addressed. E-mail: lstaudt@mail.nih.gov

This article contains supporting information online at www.pnas.org/cgi/content/full/0804295105/DCSupplemental.

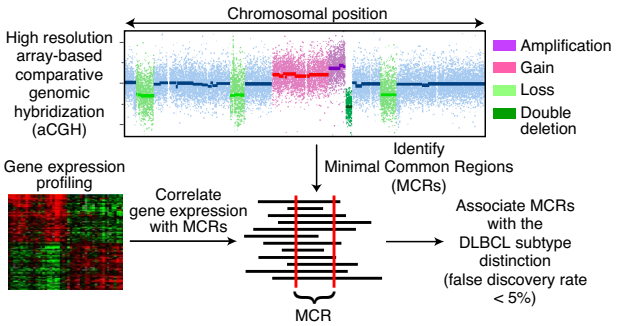

Fig. 1. Schematic of the Gene Expression and Dosage Integration algorithm (see methods).

single-copy gain, single-copy deletion, homozygous deletion, or is declared wild type [supporting information (SI) Fig. S1]. Next, GEDI defines MCRs of abnormal gene dosage by assembling segments that affect the same chromosomal region in different cases and that belong to the same dosage category (12, 13). Finally, the GEDI algorithm determines the degree to which each gene within an MCR is altered in expression and computes a statistic, summarizing these correlations across all genes in the MCR.

The GEDI algorithm yielded a total of 1940 MCRs within the DLBCL data set, ranging in size from 5 kb, affecting a single gene, to more than 240 Mb, affecting a whole chromosome. Of these MCRs, 719 showed a significant association with gene-expression alterations, whereas 1221 did not. Among the MCRs associated with gene-expression changes were 176 amplifications, 272 regions of gain or amplification (gain/amplification), 189 single-copy deletions, and 82 homozygous deletions. To focus attention on more common abnormalities, we limited subsequent analysis to the 272 MCRs that occurred in 8 or more samples (i.e., > 5% of ABC and GCB DLBCL cases) (Table S1).

Chromosomal Aberrations Associated with the DLBCL Subtype Distinction. To elucidate oncogenic mechanisms in the DLBCL subtypes, we focused our analysis on MCRs that were differentially represented among ABC DLBCLs, GCB DLBCLs, and PMBLs. We further reasoned that such MCRs might be less likely to be common copy number polymorphisms present in the human population (14).

Based on a false discovery rate (FDR) of 0.05, we observed 30 MCRs associated with the subtype distinction, far in excess of chance (Fig. S2 and Table S2). Some of these MCRs probably represent similar biology, such as single and double loss of the *INK4A/ARF* locus, both of which were much more frequent in ABC DLBCLs (Fig. 2).

Aberrations most characteristic of ABC DLBCL included trisomy 3, deletion of chromosome arm 6q, gain/amplification of a region on chromosome arm 18q, deletion of the *INK4a/ARF* tumor suppressor locus on chromosome 9, and gain/amplification of a 9-Mb region on chromosome 19. Aberrations preferentially used by GCB DLBCLs included amplification of the mir-17–92 microRNA in the *MIHG1* locus cluster on chromosome 13, gain/amplification of a 7.6-Mb region on chromosome 12, deletion of the *PTEN* tumor suppressor gene on chromosome 10, and amplification of the *REL* locus on chromosome 2. The most frequent chromosomal lesions in PMBL included amplification of a telomeric region of chromosome 9p, monosomy 10, and gain/amplification of chromosome arm 20p. Because each of these MCRs was statistically associated with changes in gene expression, it is highly likely that copy number estimates from the aCGH data were accurate. Nevertheless, we experimentally

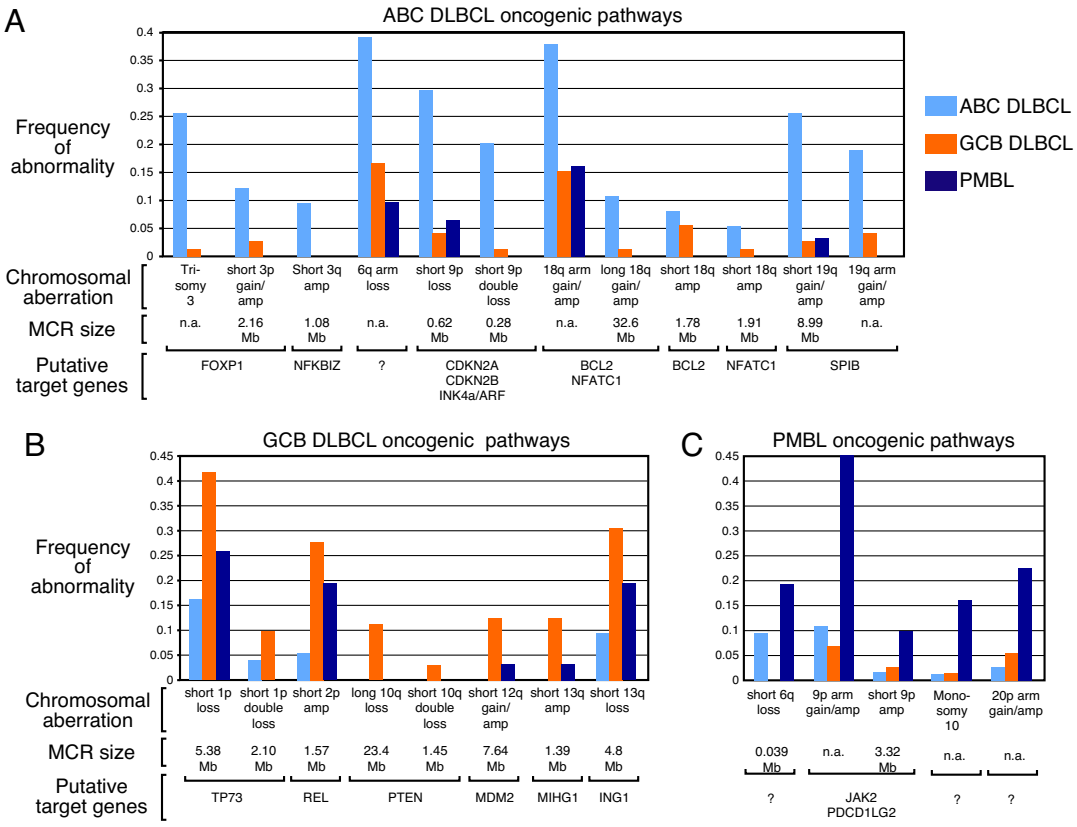

Fig. 2. Identification of DLCBL subtype-specific genomic aberrations by aCGH.

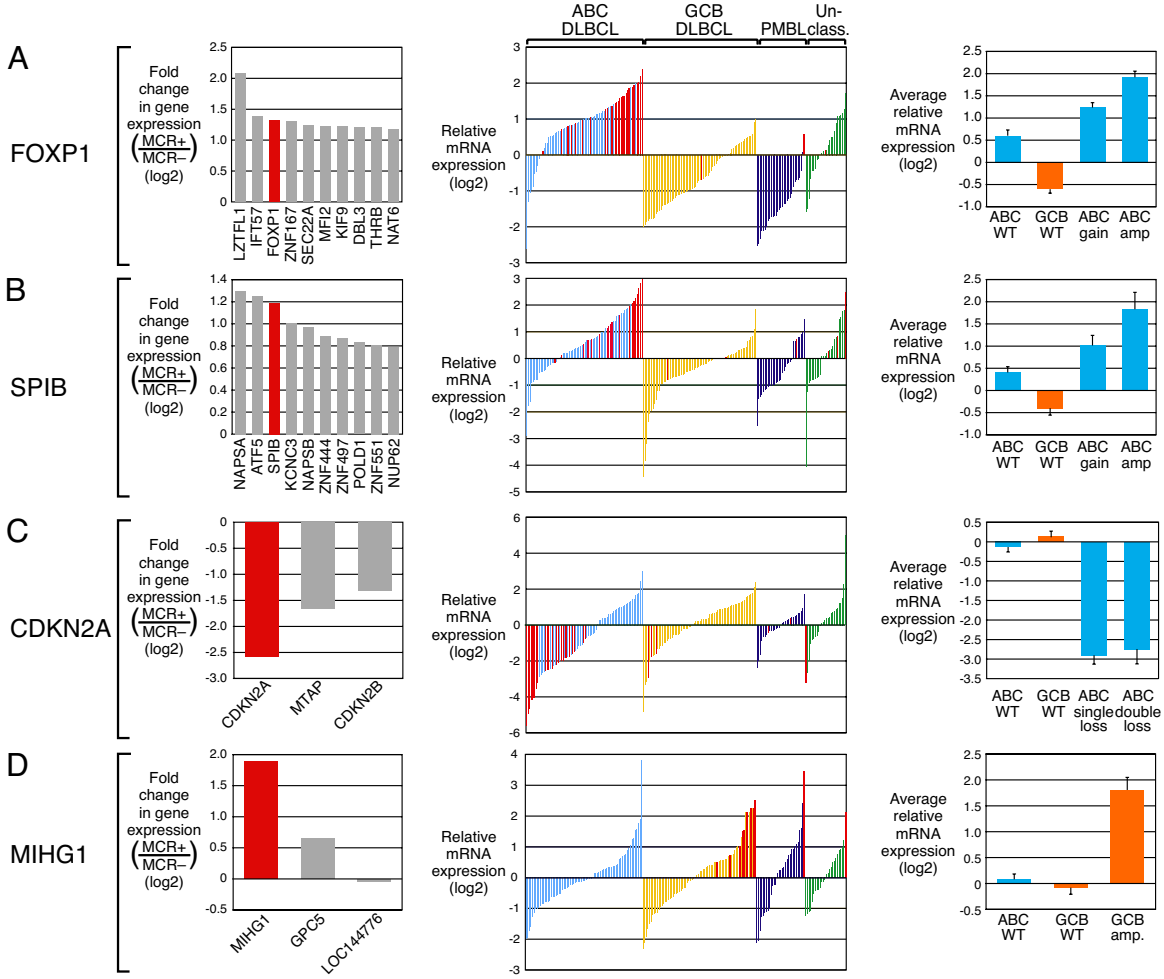

Fig. 3. Candidate oncogenes and tumor suppressors in DLBCL identified by aCGH. The left panels illustrate genes within each MCR that had the greatest fold change in cases with the aberration versus wild-type cases within the DLBCL subtype linked to each MCR. The middle panels show the expression levels of the candidate gene in each case, with red bars indicating cases with the aberration. The right panels show the average expression of the candidate gene in cases grouped as indicated. (*A*) *FOXP1* is a candidate oncogene associated with trisomy 3 in ABC DLBCL. (*B*) *SPIB* is a candidate oncogene associated with a gain/amp MCR on chromosome 19 in ABC DLBCL. (*C*) *CDKN2A* (p16) is a candidate tumor suppressor associated with single/double deletion on chromosome 9. *CDKN2B* (p15) and $p14^{ARF}$ are other tumor suppressors encoded in this MCR. (*D*) *MIHG1*, encoding the mir-17–92 microRNA cluster, is a candidate oncogene associated with an amplification in GCB DBLCL.

verified the *INK4a/ARF* locus deletion and the chromosome 19 gain/amplification in ABC DLBCL samples by quantitative PCR (Fig. S3 and Table S4).

Putative Oncogenes and Tumor Suppressors. Key to our analysis was the ability to correlate gene expression with copy number changes globally, allowing us to identify putative oncogenes and tumor suppressors affected by a change in copy number.

ABC DLBCL. Trisomy 3 was a frequent aberration in ABC DLBCL (26%) but was infrequent in GCB DLBCL (1%) and never was observed in PMBL (Fig. 2*A*). Of the ~1092 annotated genes on chromosome 3, *FOXP1* was the third most strongly up-regulated gene in cases with trisomy 3 (Fig. 3*A*); this finding was noteworthy because high *FOXP1* mRNA expression is a hallmark of ABC DLBCL (5, 15). *FOXP1* mRNA levels also were elevated in ABC DLBCL by gain of chromosome arm 3p but not 3q, gain of a 2.16-Mb region encompassing *FOXP1*, and high-level amplifications of *FOXP1* (Figs. 2*A* and 3*A*). Altogether, 38% of ABC DLBCLs had an increased *FOXP1* copy number, whereas this increase occurred in only 4% of GCB DLBCLs and 3% of PMBLs. ABC DLBCLs with *FOXP1* gain or *FOXP1* amplification had, respectively, 3.4-fold and 5.7-fold higher *FOXP1* mRNA expression than GCB DLBCLs (Fig. 3*A*). However, ABC DLBCL cases lacking these aberrations also had 2.3-fold higher *FOXP1* expression than GCB DLBCLs. Thus, *FOXP1* is transcriptionally activated in most ABC DLBCLs and is further up-regulated by either trisomy 3 or by focal *FOXP1* gain or amplification. *FOXP1* has been implicated as an oncogene based on rare translocations in DLBCL and gastric MALT lymphoma (16, 17), but trisomy 3 may be another mechanism to deregulate this gene.

Another characteristic lesion in ABC DLBCL was a 1.08-Mb amplicon on chromosome 3 that was present in 9% of cases but never was detected in GCB DLBCLs or PMBLs. This amplification includes *NFKBIZ*, which encodes an IκB-like protein that binds to NF-κB heterodimers and enhances transactivation of some NF-κB targets, such as IL-6, while repressing others (18). *NFKBIZ* is a particularly intriguing target gene, given the constitutive activation of the NF-κB pathway in ABC DLBCL (19–21) and the important role of IL-6 signaling through STAT3 in a subset of ABC DLBCLs (22).

Gain or amplification of the entire 18q chromosome arm or of a 33-Mb telomeric region of 18q occurred in 38% and 11% of

ABC DLBCLs, respectively, but these events were significantly less frequent in the other subtypes (Fig. 2). The 2 most up-regulated genes in the telomeric MCR were *BCL2* and *NFATC1*, both of which also were significantly up-regulated by gain/amplification of the 18q arm (Fig. S4). *BCL2* and *NFATC1* also were the most overexpressed genes in 2 shorter amplicons spanning 1.8 Mb and 1.9 Mb, respectively. Because these focal amplicons also were more common in ABC DLBCL than in the other subtypes, it seems likely that both BCL2 and NFATC1 contribute to the pathogenesis of ABC DLBCL.

The *INK4a/ARF* tumor suppressor locus was homozygously or hemizygously deleted in 30% of ABC DLBCLs but in only 4% of GCB DLBCLs and in 6% of PMBLs (Fig. 2). The MCRs for single and double loss of this locus are both short (620 kB and 284 kB, respectively) and encode 3 tumor suppressors, CDKN2A (p16), CDKN2B (p15), and $p14^{ARF}$. ABC DLBCL cases with hemizygous and homozygous loss of this locus had, respectively, 6.9-fold and 6.2-fold lower levels of *CDKN2A* mRNA than cases without these abnormalities (Fig. 3*C*). This finding suggests that cases with hemizygous loss have additional genetic or epigenetic processes that silence *CDKN2A* expression.

ABC DLBCLs had frequent aberrations of chromosome arm 19q, either gain/amplification of the entire arm (19%) or of an 8.99-Mb telomeric segment (26%), events that were uncommon in GCB DLBCL and PMBL (Fig. 2*A*). The telomeric MCR is very gene rich, encompassing 333 genes, but inspection of the 10 most overexpressed genes (Fig. 3*B*) revealed that 5 (*SPIB*, *NAPSA*, *NAPSB*, *POLD1*, and *KCNC3*) reside within a relatively short 113-kB interval, raising the possibility that they share a common regulatory element(s). Of these, *SPIB* appeared most likely to be functionally important because recent work uncovered a translocation between *SPIB* and the Ig heavy-chain locus in an ABC DLBCL cell line (23). *SPIB* encodes an ETS family transcription factor expressed in lymphocytes and plasmacytoid dendritic cells that is required for normal germinal center reactions (24). In ABC DLBCL, gain or amplification of the *SPIB* locus increased *SPIB* mRNA expression by 1.8-fold compared with wild-type cases, but *SPIB* copy number gain did not increase expression of *SPIB* in GCB DLBCLs. Of note, *SPIB* expression in ABC DLBCL cases with wild-type *SPIB* copy number was 1.7-fold higher than in GCB DLBCLs ($P = 3.5 \times 10^{-4}$). Thus, *SPIB* expression is characteristically elevated in ABC DLBCL and is further up-regulated by recurrent copy number gains in this subtype.

To examine the functional significance of *SPIB* in ABC DLBCL, we knocked down its expression using retroviruses expressing shRNAs, which mediate RNA interference (Fig. 4*A*) (21). Two different *SPIB* shRNAs decreased expression of *SPIB* mRNA significantly (Fig. 4*B* and Fig. S5) and were toxic to 3 ABC DLBCL cell lines (OCI-Ly3, OCI-Ly10, HBL-1) (Fig. 4*A*), each of which bears the chromosome 19 telomeric MCR containing SPIB (*data not shown*). By contrast, these 2 *SPIB* shRNAs had little or no toxicity for GCB DLBCL, PMBL, or multiple myeloma cell lines (Fig. 4*A*). A control shRNA targeting *MYC* was toxic to all cell lines (*data not shown*). These results implicate SPIB as an oncogene that is recurrently dysregulated by chromosomal abnormalities in ABC DLBCL.

GCB DLBCL. A 1.4-Mb amplicon on chromosome 13 was detected frequently in GCB DLBCL (12.5%), rarely in PMBL (3%), and never in ABC DLBCL (Fig. 2*B*). Of the 3 genes in this MCR, the most overexpressed was *MIHG1*, which produces a non-coding RNA that harbors the mir-17–92 microRNA polycistronic cluster. This amplicon has been identified previously in DLBCL, but its restriction to GCB DLBCL was not noted (25). The mir-17–92 microRNA cluster cooperates with MYC to transform mouse B cells, and lymphomas that develop with mir-17–92 overexpression have less apoptosis than those expressing MYC alone (26).

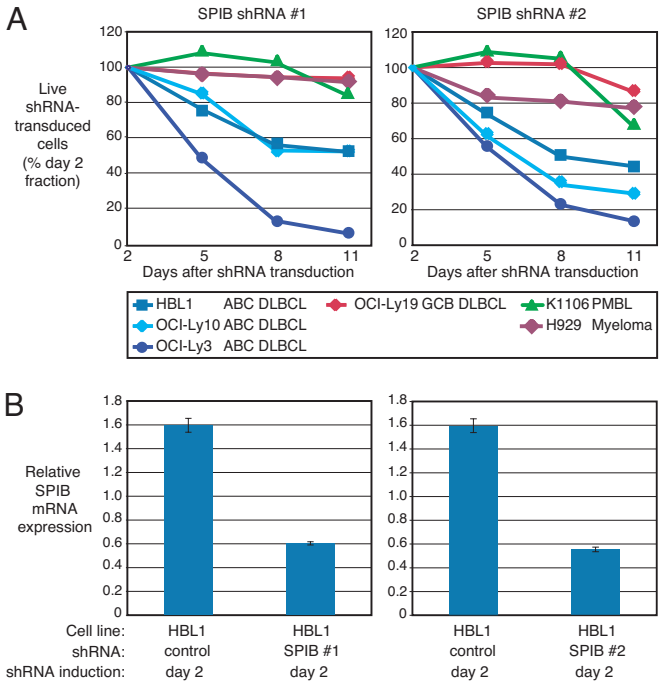

Fig. 4. SPIB is required for survival of ABC DLBCL cells. (*A*) Two independent SPIB shRNAs were toxic to ABC DLBCL cell lines but not to GCB DLBCL, PMBL, and myeloma cell lines. (*B*) Quantitative PCR analysis of *SPIB* mRNA 48 h after induction of SPIB shRNAs.

Interestingly, cases with amplification and overexpression of the mir17–92 microRNA cluster expressed *MYC* and MYC target genes at significantly higher levels than cases without this abnormality (Fig. S6). Why ABC DLBCLs lack this amplicon is unclear, but it is conceivable that constitutive activity of the anti-apoptotic NF-κB pathway in ABC DLBCLs (19, 21) eliminates the selective pressure for the anti-apoptotic action of the mir-17–92 microRNA cluster.

As noted previously (4), a locus on chromosome 2p is amplified recurrently in GCB DLBCL and PMBL, an event that is rare in ABC DLBCL (Fig. 2*B*). Among the most up-regulated genes associated with this MCR is *REL*, which encodes an NF-κB transcription factor subunit. Whereas the NF-κB pathway is activated in ABC DLBCLs by cell autonomous mechanisms (19, 21), GCB DLBCLs may use *REL* amplification to enhance the NF-κB response to signals from the microenvironment. Alternatively, *PUS10*, a gene encoding a pseudouridine kinase that is 223 kB from *REL*, may be functionally important because it is the most significantly up-regulated gene in this MCR.

Several MCRs implicate tumor suppressor pathways that may contribute to the pathogenesis of GCB DLBCL. An MCR containing *PTEN* was lost in 11% of GCB DLBCL but never in ABC DLBCL or PMBL. Moreover, homozygous deletion of a shorter 1.45-Mb region containing *PTEN* was detected in 2 GCB DLBCLs but not in other lymphoma subtypes (Fig. 2*B*). Of note, 4 of 5 GCB DLBCLs with *PTEN* deletion also had a t(14, 18) translocation, in keeping with a previous report (27), suggesting that *PTEN* inactivation and AKT pathway activation may contribute to the pathogenesis of GCB DLBCL with t(14, 18). Other MCRs may alter the response of GCB DLBCLs to DNA damage. Among the most up-regulated genes in a 7.64-Mb gain/amplification MCR on chromosome 12 was *MDM2*, encoding a negative regulator of the tumor suppressor p53. A p53 family member, p73, is a candidate tumor suppressor on chromosome 1 involved in a 5.4-Mb single-loss MCR and a 2.1-Mb homozygous deletion MCR, both of which occurred most frequently in

GENETICS

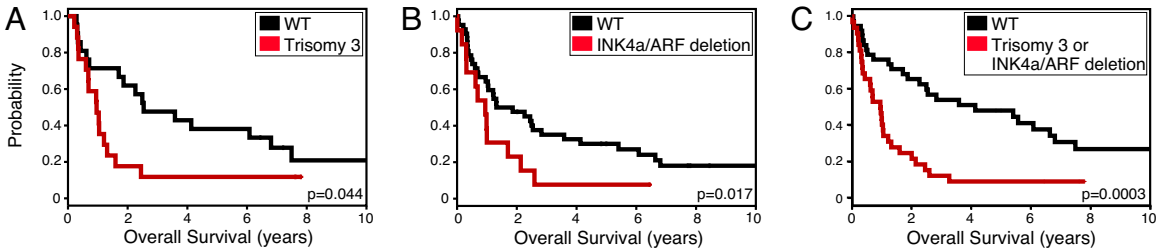

Fig. 5. Genomic aberrations associated with survival in ABC DLBCL. (*A*) Trisomy 3 and (*B*) single/double deletion of the *INK4a/ARF* locus were associated with adverse survival in ABC DLBCL. (*C*) Cases with either trisomy 3 or single/double deletion of the *INK4a/ARF* locus had inferior overall survival within ABC DBLCL.

GCB DLBCL but also occurred in the other subtypes. Finally, a 4.8-Mb single-loss MCR on chromosome 13 encoding the tumor suppressor *ING1* was observed in 31% of GCB DLBCLs but in only 9.5% of ABC DLBCLs. ING1 is a growth-suppressing protein whose action may be mediated by its interactions with p53, PCNA, or methylated histone (28).

PMBL. PMBLs had frequent amplification of 9p24, which was detected in 45% of PMBLs but in only 11% of ABC DLBCLs and in 7% of GCB DLBCLs. This MCR has been reported previously to occur in both PMBL and Hodgkin's lymphoma (6, 29), 2 lymphomas that share a common gene-expression signature, clinical presentation in younger patients, and possible origin from thymic B cells (6, 7). It is unclear which of the many genes in this MCR contribute to the pathogenesis of PMBL and Hodgkin's lymphoma, but the genes encoding the tyrosine kinase JAK2 and the T-cell inhibitory ligand PD-L2 were among the most strongly up-regulated in expression.

Genomic Loci with Prognostic Significance. Finally we investigated the association of MCRs with overall survival in our DLBCL cohort. At a stringent FDR of 0.01, we identified 8 MCRs that were associated significantly with survival within the entire cohort ($P < 0.0002$) (Tables S3 and S4). Particularly noteworthy were 2 MCRs that were strongly restricted to ABC DLBCLs and that predicted adverse survival within this subtype: trisomy 3 ($n = 19, P = 0.044$) (Fig. 5*A*) and *INK4a/ARF* locus single/double deletion ($n = 22$, $P = 0.017$) (Fig. 5*B*). We created a model in which any case with 1 or both of these abnormalities was declared aberrant ($n = 35$), and all other ABC DLBCLs were considered non-aberrant ($n = 39$). The aberrant cases had a distinctly inferior outcome, with a 5-year survival rate of only 9% compared with 48% for the non-aberrant cases ($P = 0.0003$) (Fig. 5*C*). Overall, these results suggest that this constellation of chromosomal aberrations may define a subset of ABC DLBCLs with an inferior prognosis.

Conclusions

Distinct cancer subtypes are associated with stereotypical genetic alterations (30), suggesting that the cell of origin of a cancer dictates, in part, the pathways that must be altered mutationally to achieve malignancy. By integrating genome-wide gene expression and copy number data sets, we have developed a high-resolution view of the genetic pathways that give rise to DLBCL. A primary objective of this analysis was to ascertain whether the ABC, GCB, and PMBL subtypes of DLBCL use distinct oncogenic pathways. The broad gene-expression signatures that distinguish these subtypes indicate that they probably arise from B cells at different stages of differentiation (3, 5, 8). Moreover, the distinction between ABC and GCB DLBCL has been associated with the response to chemotherapy in several large cohorts of patients (3, 4, 31). A large and diverse set of chromosomal copy number aberrations, far in excess of the number expected by chance, were associated primarily or exclusively with specific DLBCL subtypes. These genetic associations provide compelling evidence that the DLBCL subtypes represent discrete diseases that arise by distinct pathogenetic pathways.

By integrating gene-expression changes with genomic aberrations, we identified presumptive oncogenes and tumor suppressors that were deregulated in a DLBCL subtype-restricted fashion. *SPIB* emerged from this analysis as an oncogene that was up-regulated by recurrent chromosomal gains and amplifications in ABC DLBCL, as well as by a documented chromosomal translocation in an ABC DLBCL cell line (23). Using RNA interference, we demonstrated that SPIB is critical for the survival of ABC DLBCL cell lines but not for the survival of GCB DLBCL, PMBL, or myeloma cell lines. The characteristically high expression of *SPIB* in ABC DLBCL, even in the absence of a *SPIB* copy number increase, suggests that SPIB is a central regulator of the ABC DLBCL transcriptional program, a function that may be the rationale for its alteration by diverse genetic events in this DLBCL subtype.

Our analysis provides a rich starting point for future investigations into the molecular pathogenesis of DLBCL. We have highlighted some of the candidate oncogenes and tumor suppressors that are deregulated by copy number aberrations in DLBCL, each of which needs to be validated experimentally. Our integrated approach also might identify genes that are functional targets of large chromosomal aberrations. For example, trisomy 3 may be a frequent mechanism to deregulate FOXP1, which also can be overexpressed in association with focal amplifications and rare translocations in DLBCL (16, 32).

Finally, it is instructive to integrate the present analysis with the recent appreciation that constitutive activity of the anti-apoptotic NF-κB signaling pathway in ABC DLBCL critically involves the signaling scaffolding protein CARD11 (21). In roughly 10% of ABC DLBCLs, oncogenic activation of CARD11 is caused by somatic mutations in its coiled-coil domain (33). In the present cohort of ABC DLBCL cases, 7 had CARD11 mutations, and 5 of these cases also had trisomy 3 ($P = 0.012$; Fisher's exact test), suggesting that these oncogenic events lie on a common pathway of malignant transformation. As we integrate various forms of genomic analyses, the complexity of human cancer is finally coming into molecular clarity.

Materials and Methods

Patient Samples. Tumor biopsy specimens were obtained before treatment from 203 patients with *de novo* DLBCL, described previously (4). All samples were studied according to a protocol approved by the National Cancer Institute Institutional Review Board.

Array CGH. Genomic DNA from patient samples was extracted with the DNeasy Tissue kit (Qiagen). Normal male genomic DNA was used as reference. The aCGH protocol from NimbleGen Systems was followed. Briefly, tumor and reference DNAs were fragmented by sonication and random-prime labeled with Cy3 and Cy5 dyes, respectively. Labeled material was co-hybridized to microarrays consisting of 386,165 oligonucleotide probes spaced at ~ 5-kB intervals throughout the human genome, washed, and scanned (Axon Instru-

ments). The primary aCGH and gene expression data are available from the Gene Expression Omnibus of the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/geo) through GEO accession number GSE11318.

Gene-Expression Profiling. RNA was extracted from the biopsy specimens as described previously (3) and was profiled for gene expression using Affymetrix U133 plus 2.0 arrays. In Fig. 3, the following Affymetrix probe sets were used: FOXP1 (224837_at), SPIB (232739_at), CDKN2A (207039_at), and MIHG1 (232291_at).

Analysis of Array CGH Data. The GEDI algorithm and detailed statistical methods are presented in *SI Methods*. Briefly, each chromosome in each sample was divided into segments comprised of microarray probes with similar log ratios using the DNAcopy algorithm (www.bioconductor.org), and each segment was categorized as single-copy or homozygous deletion, single-copy gain, amplification, or wild type. MCRs were identified essentially as described (12, 13) and were refined further by including only those that were associated with gene expression changes ($P < 0.1$), according to a permutation test. The association of each MCR with the DLBCL subtype distinction was assessed using the Fisher's exact test, and a FDR estimation was developed to address the problem of multiple hypothesis testing arising from the large number of MCRs.

shRNA-Mediated RNA Interference. Cell lines were engineered to express the murine ecotropic retroviral receptor for efficient retroviral transductions and the bacterial tetracycline repressor for doxycycline-inducible shRNA expression, as described (21). shRNA-mediated RNA interference was performed as described (21). The targeting sequence of SPIB shRNAs 1 and 2 were CAAGGTTCCCTCTTGTCAGAT and GGACCTATGGACCACTATACT, respectively. Relative abundance of *SPIB* and *B2M* mRNA was determined after shRNA induction using Taqman quantitative RT-PCR kits (Applied Biosystems).

ACKNOWLEDGMENTS. This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), the National Cancer Institute (NCI), the Center for Cancer Research, and National Cancer Institute Strategic Partnering to Evaluate Cancer Signatures Grant UO1-CA 114778. G.L also was supported by a research grant from the German Research Foundation. This study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health (http://biowulf.nih.gov).

1. The Non-Hodgkin's Lymphoma Classification Project (1997) A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. *Blood* 89:3909–3918.
2. Coiffier B (2001) Diffuse large cell lymphoma. *Curr Opin Oncol* 13:325–334.
3. Alizadeh AA, *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–511.
4. Rosenwald A, *et al.* (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 346:1937–1947.
5. Wright G, *et al.* (2003) A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci USA* 100:9991–9996.
6. Rosenwald A, *et al.* (2003) Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J Exp Med* 198:851–862.
7. Savage KJ, *et al.* (2003) The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. *Blood* 102:3871–3879.
8. Staudt LM, Dave S (2005) The biology of human lymphoid malignancies revealed by gene expression profiling. *Adv Immunol* 87:163–208.
9. Bea S, *et al.* (2005) Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. *Blood* 106:3183–3190.
10. Tagawa H, *et al.* (2005) Comparison of genome profiles for identification of distinct subgroups of diffuse large B-cell lymphoma. *Blood* 106:1770–1777.
11. Chen W, *et al.* (2006) Array comparative genomic hybridization reveals genomic copy number changes associated with outcome in diffuse large B-cell lymphomas. *Blood* 107:2477–2485.
12. Aguirre AJ, *et al.* (2004) High-resolution characterization of the pancreatic adenocarcinoma genome. *Proc Natl Acad Sci USA* 101:9067–9072.
13. Tonon G, *et al.* (2005) High-resolution genomic profiles of human lung cancer. *Proc Natl Acad Sci USA* 102:9625–9630.
14. Redon R, *et al.* (2006) Global variation in copy number in the human genome. *Nature* 444:444–454.
15. Shaffer AL, Rosenwald A, Staudt LM (2002) Lymphoid malignancies: The dark side of B-cell differentiation. *Nat Rev Immunol* 2:920–932.
16. Wlodarska I, *et al.* (2005) FOXP1, a gene highly expressed in a subset of diffuse large B-cell lymphoma, is recurrently targeted by genomic aberrations. *Leukemia* 19:1299–1305.
17. Streubel B, Vinatzer U, Lamprecht A, Raderer M, Chott A (2005) t(3;14)(p14.1;q32) Involving IGH and FOXP1 is a novel recurrent chromosomal aberration in MALT lymphoma. *Leukemia* 19:652–658.
18. Motoyama M, Yamazaki S, Eto-Kimura A, Takeshige K, Muta T (2005) Positive and negative regulation of nuclear factor-kappaB-mediated transcription by IkappaB-zeta, an inducible nuclear protein. *J Biol Chem* 280:7444–7451.
19. Davis RE, Brown KD, Siebenlist U, Staudt LM (2001) Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J Exp Med* 194:1861–1874.
20. Lam LT, *et al.* (2005) Small molecule inhibitors of IkB-kinase are selectively toxic for subgroups of diffuse large B cell lymphoma defined by gene expression profiling. *Clin Cancer Res* 11:28–40.
21. Ngo VN, *et al.* (2006) A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 441:106–110.
22. Lam LT, *et al.* (2008) Cooperative signaling through the signal transducer and activator of transcription 3 and nuclear factor-{kappa}B pathways in subtypes of diffuse large B-cell lymphoma. *Blood* 111:3701–3713.
23. Lenz G, *et al.* (2007) Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. *J Exp Med* 204:633–643.
24. Su GH, *et al.* (1997) Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *EMBO J* 16:7118–7129.
25. Ota A, *et al.* (2004) Identification and characterization of a novel gene, C13orf25, as a target for 13q31–q32 amplification in malignant lymphoma. *Cancer Res* 64:3087–3095.
26. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT (2005) c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435:839–843.
27. Siebert R, *et al.* (1998) Deletions in the long arm of chromosome 10 in lymphomas with t(14;18): A pathogenetic role of the tumor suppressor genes PTEN/MMAC1 and MXI1? *Blood* 92:4487–4489.
28. Soliman MA, Riabowol K (2007) After a decade of study-ING, a PHD for a versatile family of proteins. *Trends Biochem Sci* 32:509–519.
29. Joos S, *et al.* (2000) Genomic imbalances including amplification of the tyrosine kinase gene JAK2 in CD30+ Hodgkin cells. *Cancer Res* 60:549–552.
30. Thomas RK, *et al.* (2007) High-throughput oncogene mutation profiling in human cancer. *Nat Genet* 39:347–351.
31. Hummel M, *et al.* (2006) A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med* 354:2419–2430.
32. Haralambieva E, *et al.* (2006) Genetic rearrangement of FOXP1 is predominantly detected in a subset of diffuse large B-cell lymphomas with extranodal presentation. *Leukemia* 20:1300–1303.
33. Lenz G, *et al.* (2008) Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* 319:1676–1679.